

# Cumulants on Wiener chaos: moderate deviations and the fourth moment theorem

Matthias Schulte\* and Christoph Thäle†

## Abstract

A moderate deviation principle as well as moderate and large deviation inequalities for a sequence of elements living inside a fixed Wiener chaos associated with an isonormal Gaussian process are shown. The conditions under which the results are derived coincide with those of the celebrated fourth moment theorem of Nualart and Peccati. The proofs rely on sharp estimates for cumulants. As applications, explosive integrals of a Brownian sheet, a discretized version of the quadratic variation of a fractional Brownian motion and the sample bispectrum of a spherical Gaussian random field are considered.

**Keywords.** Contractions, cumulants, isonormal Gaussian process, large deviation probabilities, matching number, moderate deviations, multiple stochastic integral, Wiener chaos.

**MSC.** Primary 60F10, 60H05; Secondary 60G15.

## 1 Introduction

In the last few years, the theory around probabilistic approximations of multiple stochastic integrals has advanced significantly. One of the cornerstones in this context is the so-called fourth moment theorem of Nualart and Peccati. To state it, let for simplicity  $(A, \mathcal{A})$  be a Polish space equipped with a non-atomic  $\sigma$ -finite measure  $\mu$  and let  $(h_n : n \geq 1)$  be a sequence of symmetric, measurable and square-integrable functions on  $A^q$  for a fixed integer  $q \geq 2$ . We assume that  $h_n$  satisfies  $\|h_n\|_{L^2(A^q)} = 1$  for all  $n \geq 1$ , where  $\|\cdot\|_{L^2(A^q)}$  stands for the usual norm on  $L^2(A^q)$ . Now, denote by  $F_n = I_q(h_n)$  the multiple stochastic integral of order  $q$  of  $h_n$  with respect to a Gaussian random measure on  $A$  with control  $\mu$ . The case  $q = 1$  is trivial, because then  $F_n$  is a standard Gaussian random variable, and, thus, excluded. The fourth moment theorem (see [17] and also [15, Theorem 5.2.7]) asserts that the sequence  $(F_n : n \geq 1)$  converges in distribution to a Gaussian random variable with variance  $q!$  if and only if, as  $n \rightarrow \infty$ , the fourth cumulant of  $F_n$  tends to zero, i.e.,

$$\lim_{n \rightarrow \infty} \text{cum}_4(F_n) = \lim_{n \rightarrow \infty} \mathbb{E}[F_n^4] - 3(q!)^2 = 0,$$

or equivalently, if

$$\lim_{n \rightarrow \infty} \|h_n \otimes_r h_n\|_{L^2(A^{2(q-r)})} = 0 \quad \text{for all } r \in \{1, \dots, q-1\}. \quad (1)$$

Here,  $\otimes_r$  stands for the  $r$ th contraction operator, details and precise definitions follow below. In addition, one has the following estimate for the rate of convergence (see [13] or combine Equations (5.2.6) and (5.2.13) in [15]):

$$\begin{aligned} d_{TV}(F_n, N) &:= \sup_{B \subset \mathbb{R} \text{ Borel set}} |\mathbb{P}(F_n \in B) - \mathbb{P}(N \in B)| \\ &\leq c_q \max_{r=1, \dots, q-1} \|h_n \otimes_r h_n\|_{L^2(A^{2(q-r)})}, \end{aligned} \quad (2)$$

---

\*Karlsruhe Institute of Technology, Department of Mathematics, D-76128 Karlsruhe, Germany. E-mail: matthias.schulte@kit.edu

†Ruhr University Bochum, Faculty of Mathematics, D-44780 Bochum, Germany. E-mail: christoph.thaele@rub.de

where  $N$  stands for a centred Gaussian random variable with variance  $q!$  and  $c_q > 0$  is a constant only depending on  $q$ . Note that the variance  $q!$  comes from our normalization  $\|h_n\|_{L^2(A^q)} = 1$  also used below. We emphasize that in [16] an improved and in fact optimal rate for  $d_{TV}(F_n, N)$  has been derived in terms of the third and the fourth cumulant of  $F_n$ .

The aim of this paper is to show that under condition (1) the sequence  $(F_n : n \geq 1)$  also satisfies a moderate deviation principle and fulfils moderate and large deviation inequalities. This is a direct refinement of the fourth moment theorem. A key step in our proof is to control the growth of cumulants. Such an approach has previously been used in [14] to give an alternative proof of the fourth moment theorem. To extend this to moderate deviations, we derive considerably sharper estimates for cumulants by combining classical cumulant formulas for multiple stochastic integrals with lower bounds on the matching number of regular multi-graphs. We use them together with classical large deviation results of Bentkus, Rudzkis, Saulis and Statulevičius (see the book [21]) and a transfer principle from the paper [5] of Döring and Eichelsbacher, which allows to deduce moderate deviation principles from cumulant bounds. To the best of our knowledge, the moderate deviation principle for sequences of multiple stochastic integrals is new. As applications of our general results we derive moderate deviation principles and related deviation inequalities for explosive integrals of a Brownian sheet, a discretized version of the quadratic variation of a fractional Brownian motion on the unit interval and the sample bispectrum of a Gaussian random field on the two-dimensional unit sphere.

To motivate our results and to introduce some basic ideas, let us briefly consider the case of a sum  $S_n = X_1 + \dots + X_n$  of  $n$  independent and identically distributed centred random variables with variance  $\sigma^2 > 0$ . Let us assume that the random variables  $(X_i : i \geq 1)$  have finite exponential moments in that  $\mathbb{E}[e^{\lambda X_1}] < \infty$  for all  $|\lambda| \leq \Lambda$  and some  $\Lambda > 0$ . Denoting by  $\Phi_{\sigma^2}(\cdot)$  the distribution function of a centred Gaussian random variable with variance  $\sigma^2$ , the central limit theorem ensures that

$$\lim_{n \rightarrow \infty} \frac{\mathbb{P}(S_n/\sqrt{n} \geq z)}{1 - \Phi_{\sigma^2}(z)} = 1 \quad \text{and} \quad \lim_{n \rightarrow \infty} \frac{\mathbb{P}(S_n/\sqrt{n} \leq -z)}{\Phi_{\sigma^2}(-z)} = 1 \quad (3)$$

for fixed  $z \geq 0$ . In the theory of moderate deviations one is interested in the following two questions refining the central limit theorem.

1. How fast do the ratios in (3) converge to 1?
2. Under which conditions does relation (3) remain valid if  $z$  is growing with  $n$ , i.e., if  $z$  is replaced by  $a_n z$  with  $a_n \rightarrow \infty$ , as  $n \rightarrow \infty$ ?

One way to approach moderate deviations is to provide upper bounds for the ratios in (3) on a logarithmic scale. Here, one can show that there are constants  $n_0, z_0, c > 0$  only depending on the distribution of the random variable  $X_1$  such that, for all  $n \geq n_0$  and  $0 \leq z \leq z_0 \sqrt{n}$ ,

$$\left| \log \frac{\mathbb{P}(S_n/\sqrt{n} \geq z)}{1 - \Phi_{\sigma^2}(z)} \right| \leq c \frac{1 + (z/\sigma)^3}{\sqrt{n}}$$

(a similar result is also available for the ratio  $\mathbb{P}(S_n/\sqrt{n} \leq -z)/\Phi_{\sigma^2}(-z)$ ), answering thereby the first question. To handle the second problem stated above, one can follow Varadhan's ideas for large deviations and study the re-scaled logarithm of the probability  $\mathbb{P}(S_n/\sqrt{n} \geq a_n z) = \mathbb{P}(S_n/(a_n \sqrt{n}) \geq z)$ . In the present situation, it is known that

$$\lim_{n \rightarrow \infty} a_n^{-2} \log \mathbb{P}(S_n/(a_n \sqrt{n}) \in [z, \infty)) = -\frac{z^2}{2\sigma^2}, \quad (4)$$

whenever the sequence  $(a_n : n \geq 1)$  satisfies  $a_n \rightarrow \infty$  and  $a_n/\sqrt{n} \rightarrow 0$ , as  $n \rightarrow \infty$ . We remark that a similar relation also holds for the interval  $(-\infty, -z]$ . Extending (4) in a suitable way

from intervals  $B = [z, \infty)$  to arbitrary Borel sets  $B \subset \mathbb{R}$ , one arrives at the usual form of a large deviation principle with speed  $a_n^{-2}$  and rate function  $z^2/(2\sigma^2)$  (formal definitions will be given below). Since the rate function is inherited from the central limit theorem, this large deviation principle is usually referred to as a moderate deviation principle.

The rest of this paper is structured as follows. In Section 2 we introduce the notation and our general set-up. The main results are stated in Section 3, while Sections 4, 5 and 6 present the applications to Brownian sheet, fractional Brownian motion and spherical Gaussian random fields. The proofs of the main results are the content of the final Section 7.

## 2 Preliminaries

Through this paper,  $(\Omega, \mathcal{F}, \mathbb{P})$  denotes an underlying probability space and expectation with respect to  $\mathbb{P}$  is indicated by  $\mathbb{E}$ . By  $L^2(\Omega, \mathcal{F}, \mathbb{P})$  we denote the space of square-integrable random variables on  $(\Omega, \mathcal{F}, \mathbb{P})$ .

**Large and moderate deviation principles.** Let  $(X_n : n \geq 1)$  be a sequence of real-valued random variables, let  $(s_n : n \geq 1)$  be a sequence of positive real numbers such that  $s_n \rightarrow \infty$ , as  $n \rightarrow \infty$ , and let  $\mathcal{I} : \mathbb{R} \rightarrow [0, \infty]$  be a lower semi-continuous function with compact level sets such that  $\mathcal{I} \not\equiv 0$  and  $\mathcal{I} \not\equiv \infty$ . We say that the sequence  $(X_n : n \geq 1)$  satisfies a large deviation principle (LDP) with speed  $s_n$  and rate function  $\mathcal{I}$  if for all Borel sets  $B \subset \mathbb{R}$  we have that

$$\liminf_{n \rightarrow \infty} s_n^{-1} \log \mathbb{P}(X_n \in B) \geq - \inf_{z \in \text{int}(B)} \mathcal{I}(z)$$

and

$$\limsup_{n \rightarrow \infty} s_n^{-1} \log \mathbb{P}(X_n \in B) \leq - \inf_{z \in \text{cl}(B)} \mathcal{I}(z),$$

where  $\text{int}(B)$  and  $\text{cl}(B)$  stand for the interior and the closure of  $B$ , respectively, see Chapter 1.2 in [4].

In our paper, we will deal with a special class of LDPs. To introduce them, fix a real-valued sequence  $(a_n : n \geq 1)$  such that  $a_n \rightarrow \infty$ , as  $n \rightarrow \infty$ , and suppose that the random variables  $(X_n : n \geq 1)$  satisfy a central limit theorem with limiting Gaussian distribution  $\mathcal{N}(0, \sigma^2)$  for some  $\sigma^2 > 0$ . In the following, we say that the re-scaled sequence  $(a_n^{-1}X_n : n \geq 1)$  satisfies a moderate deviation principle (MDP) if it satisfies a LDP with speed  $s_n = a_n^2$  and Gaussian rate function  $\mathcal{I}(z) = z^2/(2\sigma^2)$ . Typically, a MDP is valid for a whole range of scales  $(a_n : n \geq 1)$ .

**Isonormal Gaussian processes and chaotic representation.** In this paper we deal with sequences living inside a fixed Wiener chaos associated with an isonormal Gaussian process. To introduce the set-up formally, let here and through the rest of this paper  $\mathfrak{H}$  be a real separable Hilbert space with inner product  $\langle \cdot, \cdot \rangle_{\mathfrak{H}}$  and norm  $\| \cdot \|_{\mathfrak{H}}$ . We denote for integers  $q \geq 1$  by  $\mathfrak{H}^{\otimes q}$  the  $q$ th tensor power and by  $\mathfrak{H}^{\odot q}$  the  $q$ th symmetric tensor power of  $\mathfrak{H}$ . We supply  $\mathfrak{H}^{\otimes q}$  with the canonical scalar product  $\langle \cdot, \cdot \rangle_{\mathfrak{H}^{\otimes q}}$  and norm  $\| \cdot \|_{\mathfrak{H}^{\otimes q}}$ , while  $\mathfrak{H}^{\odot q}$  is equipped with the norm  $\sqrt{q!} \| \cdot \|_{\mathfrak{H}^{\otimes q}}$ . By  $X = (X(h) : h \in \mathfrak{H})$  we indicate an isonormal Gaussian process over  $\mathfrak{H}$  defined on our underlying probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and assume that  $\mathcal{F} = \sigma(X)$ . In other words,  $X$  is a family consisting of centred Gaussian random variables such that  $\mathbb{E}[X(h)X(h')] = \langle h, h' \rangle_{\mathfrak{H}}$  for  $h, h' \in \mathfrak{H}$ . The  $q$ th Wiener chaos  $\mathcal{C}_q$  associated with  $X$  is the closed linear subspace of  $L^2(\Omega, \mathcal{F}, \mathbb{P})$ , which is generated by random variables of the form  $H_q(X(h))$ , where  $H_q$  is the  $q$ th Hermite polynomial and  $h \in \mathfrak{H}$  satisfies  $\|h\|_{\mathfrak{H}} = 1$ . We also put  $\mathcal{C}_0 := \mathbb{R}$ . It is well known that the mapping  $h^{\otimes q} \mapsto H_q(X(h))$  can be extended to a linear isometry  $I_q$  from  $\mathfrak{H}^{\odot q}$  to  $\mathcal{C}_q$ , cf. [15, Chapter 2]. For  $q = 0$  put  $I_0(c) = c$  for all  $c \in \mathbb{R}$ . In the particular case that  $\mathfrak{H} = L^2(A)$  with a Polish space  $(A, \mathcal{A})$  and a non-atomic  $\sigma$ -finite measure  $\mu$ ,  $I_q(\cdot)$  has an interpretation as multiple

stochastic integral of order  $q$  with respect to a Gaussian random measure on  $A$  with control  $\mu$  as discussed in the introduction, cf. [15, Section 2.7.1].

It is a classical result in stochastic analysis (see [15, Theorem 2.2.4], for example) that  $L^2(\Omega, \mathcal{F}, \mathbb{P})$  can be decomposed into an infinite orthogonal sum of Wiener chaoses  $\mathcal{C}_q$ ,  $q \geq 0$ . In particular, any  $F \in L^2(\Omega, \mathcal{F}, \mathbb{P})$  can be represented as

$$F = \sum_{q=0}^{\infty} I_q(h^{(q)})$$

with  $h^{(0)} = \mathbb{E}[F]$  and uniquely determined elements  $h^{(q)} \in \mathfrak{H}^{\odot q}$ ,  $q \geq 1$ . We finally notice that random variables of the form  $I_q(h)$  for  $h \in \mathfrak{H}^{\odot q}$  satisfy  $\mathbb{E}[I_q(h)] = 0$ ,  $\mathbb{E}[I_q(h)^2] = q! \|h\|_{\mathfrak{H}^{\otimes q}}^2$  and have finite moments of all orders due the hypercontractivity property [15, Theorem 2.7.2].

**Contractions.** Let  $(e_n : 1 \leq n \leq \dim \mathfrak{H})$  if  $\dim \mathfrak{H} < \infty$  or  $(e_n : n \geq 1)$  if  $\dim \mathfrak{H} = \infty$  be a complete orthonormal system in  $\mathfrak{H}$ . For integers  $p, q \geq 1$ ,  $f \in \mathfrak{H}^{\odot p}$ ,  $g \in \mathfrak{H}^{\odot q}$  and  $r \in \{1, \dots, \min(p, q)\}$  we denote by  $f \otimes_r g$  the  $r$ th contraction of  $f$  and  $g$  defined as

$$f \otimes_r g := \sum_{i_1, \dots, i_r=1}^{\dim \mathfrak{H}} \langle f, e_{i_1} \otimes \dots \otimes e_{i_r} \rangle_{\mathfrak{H}^{\otimes r}} \otimes \langle g, e_{i_1} \otimes \dots \otimes e_{i_r} \rangle_{\mathfrak{H}^{\otimes r}},$$

see [15, Chapter B.4]. We notice that  $f \otimes_q g = \langle f, g \rangle_{\mathfrak{H}^{\otimes q}}$  if  $p = q$ . In case that  $\mathfrak{H} = L^2(A)$  for a Polish space  $(A, \mathcal{A})$  and a non-atomic  $\sigma$ -finite measure  $\mu$  we have that  $\mathfrak{H}^{\odot q} = L_s^2(A^q)$  (that is the subspace of  $\mu^q$ -a.e. symmetric functions in  $L^2(A^q)$ ) and that

$$f \otimes_r g(a_1, \dots, a_{p+q-2r}) = \int_{A^r} f(x_1, \dots, x_r, a_1, \dots, a_{p-r}) \times g(x_1, \dots, x_r, a_{p-r+1}, \dots, a_{p+q-2r}) \mu^r(dx_1, \dots, x_r)$$

for  $f \in L_s^2(A^p)$ ,  $g \in L_s^2(A^q)$  and  $r \in \{1, \dots, \min(p, q)\}$ . In other words  $f \otimes_r g$  is a function of  $p + q - 2r$  arguments, which arises from the tensor product of  $f$  and  $g$  by identifying  $r$  variables which are then integrated out.

**Cumulants.** Let  $F$  be a real-valued random variable such that  $\mathbb{E}[|F|^m] < \infty$  for some integer  $m \geq 1$ . By  $\phi_F(t) := \mathbb{E}[e^{itF}]$ ,  $t \in \mathbb{R}$ , we denote the characteristic function of  $F$ , where  $i$  is the imaginary unit. Then the  $m$ th cumulant of  $F$  (sometimes also called semi-invariant) is defined as

$$\text{cum}_m(F) := (-i)^m \frac{d^m}{dt^m} \log \phi_F(t) \Big|_{t=0}. \quad (5)$$

For example,  $\text{cum}_1(F)$  coincides with the mean of  $F$ , while  $\text{cum}_2(F)$  is its variance. Moreover, for centred random variables  $F$  we have the relations  $\text{cum}_3(F) = \mathbb{E}[F^3]$  and  $\text{cum}_4(F) = \mathbb{E}[F^4] - 3(\mathbb{E}[F^2])^2$ .

### 3 Main results

Let  $\mathfrak{H}$  be a real separable Hilbert space underlying our isonormal Gaussian process and  $(h_n : n \geq 1)$  be a sequence of elements of  $\mathfrak{H}^{\odot q}$  for some fixed integer  $q \geq 2$ . To simplify some of our arguments below, we assume without loss of generality that  $\|h_n\|_{\mathfrak{H}^{\otimes q}} = 1$  for each  $n \geq 1$ . This implies that  $\text{Var } I_q(h_n) = \mathbb{E}[I_q(h_n)^2] = q! \|h_n\|_{\mathfrak{H}^{\otimes q}}^2 = q!$  for all  $n \geq 1$ . We further define, for  $n \geq 1$ ,

$$K_n := \max_{r=1, \dots, q-1} \|h_n \otimes_r h_n\|_{\mathfrak{H}^{\otimes 2(q-r)}} \quad (6)$$

and put

$$\alpha(q) := \frac{q+2}{3q+2} \quad (q \text{ even}), \quad \alpha(q) := \frac{q^2 - q - 1}{q(3q-5)} \quad (q \text{ odd}). \quad (7)$$

Our first result delivers a moderate deviation principle (MDP) for the sequence  $(I_q(h_n) : n \geq 1)$  in the regime in which  $K_n \rightarrow 0$ , as  $n \rightarrow \infty$ . As discussed in the introduction, this is precisely the situation under which the fourth moment theorem ensures that the sequence  $(I_q(h_n) : n \geq 1)$  satisfies a central limit theorem, recall (2).

**Theorem 1.** *Define  $F_n := I_q(h_n)$  and suppose that  $\lim_{n \rightarrow \infty} K_n = 0$ . Further, let  $(a_n : n \geq 1)$  be a real sequence such that*

$$\lim_{n \rightarrow \infty} a_n = \infty \quad \text{and} \quad \lim_{n \rightarrow \infty} a_n / \Delta_n^{1/(q-1)} = 0$$

*with  $\Delta_n = (q^{3q/2} K_n^{\alpha(q)})^{-1}$ . Then the sequence  $(a_n^{-1} F_n : n \geq 1)$  satisfies a MDP with speed  $a_n^2$  and Gaussian rate function  $\mathcal{I}(z) = z^2/(2q!)$ .*

Theorem 1 can be seen as a direct refinement of the fourth moment theorem of Nualart and Peccati [17]. It provides under exactly the same conditions further information about the distributional behaviour of the involved random variables in form of a MDP.

Our next result provides a version of Theorem 1, which is more amenable to some concrete applications, see Section 6 for an example. It shows that the sequence  $(F_n : n \geq 1)$  satisfies a MDP if  $\text{cum}_4(F_n) \rightarrow 0$ , as  $n \rightarrow \infty$ . Defining

$$L_n := (q q!)^{-1} \sqrt{\text{cum}_4(F_n)}, \quad n \geq 1, \quad (8)$$

we see that the following result is a direct consequence of Theorem 1 and the estimate  $K_n \leq L_n$  from [2, Equation (4.5)].

**Corollary 2.** *Let  $F_n := I_q(h_n)$  and suppose that  $\lim_{n \rightarrow \infty} \text{cum}_4(F_n) = 0$ . Then the conclusion of Theorem 1 remains valid with  $K_n$  replaced by  $L_n$  in the definition of  $\Delta_n$ .*

Note that the moderate deviation principle in Theorem 1 or Corollary 2 for the sequence of random variables  $F_n = I_q(h_n)$ ,  $n \geq 1$ , holds in a range of scales  $(a_n : n \geq 1)$ , which shrinks with growing  $q$ . The following example shows that this phenomenon is unavoidable. For this, recall that  $H_q$  stands for the  $q$ th Hermite polynomial and that the random variables of the form  $H_q(X(h)) = I_q(h^{\otimes q})$  with  $h \in \mathfrak{H}$  and  $\|h\|_{\mathfrak{H}} = 1$  are the basic building blocks of the  $q$ th Wiener chaos  $\mathcal{C}_q$ .

**Proposition 3.** *Suppose that  $\dim \mathfrak{H} = \infty$ , let  $(e_k : k \geq 1)$  be a complete orthonormal system in  $\mathfrak{H}$ , fix  $q \geq 2$  and for each  $n \geq 1$  define  $F_n := \frac{1}{\sqrt{n}} \sum_{k=1}^n H_q(X(e_k))$ . Then  $F_n \in \mathcal{C}_q$ ,  $\text{Var } F_n = q!$ ,  $K_n = 1/\sqrt{n}$  and there are constants  $C, c, z_0 > 0$  such that, for any  $n \geq 1$ ,*

$$\mathbb{P}(|F_n| \geq z) \geq C \exp(-c n^{1/q} z^{2/q}), \quad z \geq z_0. \quad (9)$$

*Moreover, if  $(a_n : n \geq 1)$  is a real sequence such that*

$$\lim_{n \rightarrow \infty} a_n = \infty \quad \text{and} \quad \lim_{n \rightarrow \infty} a_n / n^{1/(2q-2)} = \infty, \quad (10)$$

*then the sequence  $(a_n^{-1} F_n : n \geq 1)$  does not satisfy a MDP with speed  $a_n^2$  and Gaussian rate function  $\mathcal{I}(z) = z^2/(2q!)$ .*

The random variables  $(F_n : n \geq 1)$  considered in Proposition 3 satisfy the assumptions of Theorem 1. It asserts that for a real sequence  $(a_n : n \geq 1)$  the random variables  $(a_n^{-1} F_n : n \geq 1)$  satisfy a MDP with speed  $a_n^2$  and Gaussian rate function  $\mathcal{I}(z) = z^2/(2q!)$  if

$$\lim_{n \rightarrow \infty} a_n = \infty \quad \text{and} \quad \lim_{n \rightarrow \infty} a_n / n^{\alpha(q)/(2q-2)} = 0. \quad (11)$$

Comparing conditions (10) and (11), we see that there are re-scalings  $(a_n : n \geq 1)$  such that the sequence  $(a_n^{-1}F_n : n \geq 1)$  of random variables from Proposition 3 satisfies a MDP and re-scalings  $(a_n : n \geq 1)$  for which no such MDP can hold. If the growth of  $(a_n : n \geq 1)$  is between that of  $n^{\alpha(q)/(2q-2)}$  and  $n^{1/(2q-2)}$ , it remains an open problem whether  $(a_n^{-1}F_n : n \geq 1)$  satisfies a MDP or not. We emphasize that via the method of cumulants – which is the key tool to derive Theorem 1 – it is not possible to decide this question. More precisely, our proof of Theorem 1 relies on a combination of cumulant estimates, lower bounds for the matching number of regular multigraphs and the general theory of large deviations from [21] and its recent extension [5]. We only need to bound the cumulants of  $F_n$  from above. However, since there are situations in which all of our estimates in the proof of Theorem 1 are sharp, we see that our cumulant bounds can in general not be improved as further discussed in Remark 19. This in turn implies that the range of re-scalings  $(a_n : n \geq 1)$  of the MDP in Theorem 1 or Corollary 2 is the best one can achieve by the method of cumulants. We remark that cumulant estimates for multiple stochastic integrals have also been performed in the proof of Theorem 5.8 in [14]. While these bounds are sufficient to yield the fourth moment theorem, they are far too crude in order to show our results.

**Remark 4.** Usually a large deviation principle is called a moderate deviation principle if the magnitude of re-scaling is between that of the central limit theorem and that of a law of large numbers. For the random variables in Theorem 1 the latter one does not apply since in general there is no law of large numbers. However, for  $a_n = \sqrt{n}$  the random variables  $(a_n^{-1}F_n : n \geq 1)$  in Proposition 3 satisfy the classical strong law of large numbers, which justifies the denotation as MDP.

Our next theorem deals with moderate and large deviation probabilities. To state it, recall that  $\Phi_{\sigma^2}$  denotes the distribution function of a centred Gaussian random variable with variance  $\sigma^2$ .

**Theorem 5.** Let  $F_n := I_q(h_n)$  and put  $\Delta_n := (q^{3q/2} K_n^{\alpha(q)})^{-1}$ .

(i) There are constants  $c_0, c_1, c_2 > 0$  depending only on  $q$  such that, for  $\Delta_n \geq c_0$  and  $0 \leq z \leq c_1 \Delta_n^{1/(q-1)}$ ,

$$\left| \log \frac{\mathbb{P}(F_n \geq z)}{1 - \Phi_{q!}(z)} \right| \leq c_2 \frac{1 + (z/\sqrt{q!})^3}{\Delta_n^{1/(q-1)}} \quad \text{and} \quad \left| \log \frac{\mathbb{P}(F_n \leq -z)}{\Phi_{q!}(-z)} \right| \leq c_2 \frac{1 + (z/\sqrt{q!})^3}{\Delta_n^{1/(q-1)}}.$$

(ii) For all  $z \geq 0$  one has

$$\mathbb{P}(|F_n| \geq z) \leq 2 \exp \left( -\frac{1}{4} \min \left\{ \frac{z^2}{2^{q/2}}, (z\Delta_n)^{2/q} \right\} \right). \quad (12)$$

(iii) The statements (i) and (ii) remain valid if in the definition of  $\Delta_n$ ,  $K_n$  is replaced by  $L_n$  given by (8).

**Remark 6.** (i) The first part of Theorem 5 is a consequence of a version of the celebrated ‘lemma’ of Rudzkis, Saulis and Statulevičius [20] applied to  $F_n$ . Its original formulation involves the so-called Cramér-Petrov series. For clarity and to avoid heavy notation, we have decided to state here the result in a simplified form taken from [6, Corollary 3.2], suppressing thereby higher-order terms of the expansion.

(ii) The second part of Theorem 5 follows from a result of Bentkus and Rudzkis [1]. Our statement is a simplified version taken from the Corollary after Lemma 2.4 in [21].

(iii) As Corollary 2, Theorem 5 (iii) is again a consequence of the estimate  $K_n \leq L_n$ .

Let us relate our result obtained in Theorem 5 (ii) to the existing literature, especially to the tail estimates for multiple stochastic integrals established by Major [9]. Theorem 8.5 there says that there is a constant  $c > 0$  only depending on  $q$  such that  $F_n = I_q(h_n)$  with  $\|h_n\|_{\mathfrak{H}^{\otimes q}} = 1$  satisfies

$$\mathbb{P}(|F_n| > z) \leq c \exp\left(-\frac{1}{2}\left(\frac{z}{\sqrt{q!}}\right)^{2/q}\right) \quad \text{for all } z \geq 0. \quad (13)$$

We also refer to [7, Theorem 6.12] for a similar result. If  $z$  and  $\Delta_n$  are sufficiently large, the minimum in (12) is larger than  $(z/\sqrt{q!})^{2/q}/2$ , whence (12) yields better estimates than (13). This is the case if  $K_n$  is small, meaning in view of (2) that the distribution of  $F_n$  is close to a Gaussian distribution.

Moreover, we would like to mention that Chapter 5.2 of [21] also contains a set of large deviation inequalities for multiple stochastic integrals. These results only involve the  $L^2$ -norms of the integrands and not the contractions as encoded in the sequence  $K_n$  defined at (6) so that a connection to the central limit theorem remained hidden. Applying the Cauchy-Schwarz inequality to  $K_n$ , we can recover the results in [21] from our Theorem 5. We also point out that in [21] no connection has been made to a MDP as stated in Theorem 1.

## 4 Application to Brownian sheet

Let  $W = (W_t : t \in [0, 1])$  be a standard Brownian motion on the unit interval. Then

$$\mathbb{E} \int_0^1 \frac{W_t^2}{t^2} dt = \int_0^1 \frac{\mathbb{E} W_t^2}{t^2} dt = \int_0^1 \frac{1}{t} dt = \infty,$$

and Jeulin's Lemma (see [19]) implies that  $\int_0^1 \frac{W_t^2}{t^2} dt = \infty$  with probability one. However, for any  $n \geq 2$  the functional

$$G_n := \int_{1/n}^1 \frac{W_t^2}{t^2} dt \quad (14)$$

has mean  $\mathbb{E}[G_n] = \log n$  and is finite with probability one. It is thus natural to describe the rate and type of 'explosion' of the random integral, as  $n \rightarrow \infty$ . We emphasize that such explosive random integrals have some fundamental connections with the theory of enlargement of filtrations and to Brownian local times as further discussed in [19], see also [8].

More generally, we consider a similar family of random explosive integrals with respect to a standard Brownian sheet  $\mathbf{W} = (\mathbf{W}(t_1, \dots, t_d) : (t_1, \dots, t_d) \in [0, 1]^d)$  on the  $d$ -dimensional unit cube  $[0, 1]^d$  for some fixed space dimension  $d \geq 1$ . Recall that  $\mathbf{W}$  is a centred Gaussian random field on  $[0, 1]^d$  with covariance

$$\mathbb{E}[\mathbf{W}(t_1, \dots, t_d) \mathbf{W}(s_1, \dots, s_d)] = \prod_{i=1}^d \min(t_i, s_i), \quad t_1, \dots, t_d, s_1, \dots, s_d \in [0, 1].$$

In analogy to (14) put, for  $n \geq 2$ ,

$$G_n^{(d)} := \int_{[1/n, 1]^d} \frac{\mathbf{W}(t_1, \dots, t_d)^2}{t_1^2 \cdots t_d^2} d(t_1, \dots, t_d)$$

so that  $G_n^{(1)}$  reduces to  $G_n$ . Taking expectation and using Fubini's theorem yields that

$$\begin{aligned} \mathbb{E}[G_n^{(d)}] &= \int_{[1/n, 1]^d} \frac{\mathbb{E}[\mathbf{W}(t_1, \dots, t_d)^2]}{t_1^2 \cdots t_d^2} d(t_1, \dots, t_d) \\ &= \int_{[1/n, 1]^d} \frac{1}{t_1 \cdots t_d} d(t_1, \dots, t_d) = \left( \int_{1/n}^1 \frac{1}{t} dt \right)^d = (\log n)^d. \end{aligned}$$

As shown in Section 3.2 in [17],  $G_n^{(d)} - \mathbb{E}[G_n^{(d)}]$  can be represented as  $I_2^{\mathbf{W}}(h_n^{(d)})$  with

$$h_n^{(d)}(t_1, \dots, t_d, s_1, \dots, s_d) = \prod_{i=1}^d (\max(t_i, s_i, 1/n)^{-1} - 1),$$

where  $I_2^{\mathbf{W}}$  stands for the double stochastic integral with respect to the Brownian sheet  $\mathbf{W}$ . Hence, we have that

$$\begin{aligned} \text{Var } G_n^{(d)} &= 2 \int_{([0,1]^d)^2} h_n^{(d)}(t_1, \dots, t_d, s_1, \dots, s_d)^2 d(t_1, \dots, t_d, s_1, \dots, s_d) \\ &= 2 \left( \int_0^1 \int_0^1 (\max\{t, s, 1/n\}^{-1} - 1)^2 dt ds \right)^d \\ &= 2 \left( 2 \int_0^1 \int_0^s (\min\{1/t, 1/s, n\} - 1)^2 dt ds \right)^d \\ &= 2 \left( 2 \int_{1/n}^1 s \left( \frac{1}{s} - 1 \right)^2 ds + 2 \int_0^{1/n} s(n-1)^2 ds \right)^d \\ &= 2 \left( 2 \log n - 2 \left( 1 - \frac{1}{n} \right) \right)^d. \end{aligned}$$

We now define a normalized version of  $G_n^{(d)}$  as

$$F_n^{(d)} := \frac{G_n^{(d)} - (\log n)^d}{(2 \log n - 2(1 - 1/n))^{d/2}}, \quad n \geq 2. \quad (15)$$

This normalization ensures that  $\mathbb{E}[F_n^{(d)}] = 0$  and  $\text{Var}(F_n^{(d)}) = 2$ .

By Proposition 8 in [17] the random variables  $(F_n^{(d)} : n \geq 2)$  satisfy a central limit theorem, as  $n \rightarrow \infty$ . Our theory developed in Section 3 allows to add a moderate deviation principle as well as moderate and large deviation inequalities.

**Theorem 7.** *Let  $F_n^{(d)}$  be defined as at (15). Then the statements of Theorem 1 and Theorem 5 are valid with  $q = 2$  and  $\Delta_n = \frac{1}{8\sqrt{2}} \left( \frac{\log n}{120} \right)^{d/4}$ ,  $n \geq 2$ .*

*Proof.* In what follows, we compute an upper bound for

$$\begin{aligned} K_n &= \frac{2}{\text{Var } G_n^{(d)}} \|h_n^{(d)} \otimes_1 h_n^{(d)}\|_{L^2([0,1]^d)^2} \\ &= \frac{2}{\text{Var } G_n^{(d)}} \left( \int_0^1 \int_0^1 \left( \int_0^1 h_n(x, z) h_n(y, z) dz \right)^2 dx dy \right)^{d/2}. \end{aligned}$$

For this, we make use of the estimate

$$\int_0^1 h_n(x, z) h_n(y, z) dz \leq \int_0^1 \min(1/z, 1/x, n) \min(1/z, 1/y, n) dz =: g(x, y),$$

valid for all  $x, y \in [0, 1]$ . We now consider the cases  $1/n \leq x \leq y$ ,  $x \leq 1/n \leq y$  and  $x \leq y \leq 1/n$  separately and obtain that

$$\begin{aligned} g(x, y) &= \int_0^x \frac{1}{xy} dz + \int_x^y \frac{1}{zy} dz + \int_y^1 \frac{1}{z^2} dz = \frac{1}{y} + \frac{\log y - \log x}{y} + \frac{1}{y} - 1 \leq \frac{2}{y} + \frac{\log y - \log x}{y}, \\ g(x, y) &= \int_0^{1/n} \frac{n}{y} dz + \int_{1/n}^y \frac{1}{zy} dz + \int_y^1 \frac{1}{z^2} dz = \frac{1}{y} + \frac{\log y + \log n}{y} + \frac{1}{y} - 1 \leq \frac{2}{y} + \frac{\log y + \log n}{y}, \\ g(x, y) &= \int_0^{1/n} n^2 dz + \int_{1/n}^1 \frac{1}{z^2} dz = n + n - 1 \leq 2n, \end{aligned}$$



respectively. Together with the Cauchy-Schwarz inequality this implies that

$$\begin{aligned} \int_0^1 \int_0^1 g(x, y)^2 \, dx \, dy &\leq 2 \int_{1/n}^1 \int_{1/n}^y \frac{8}{y^2} + \frac{2(\log y - \log x)^2}{y^2} \, dx \, dy \\ &\quad + 2 \int_{1/n}^1 \int_0^{1/n} \frac{8}{y^2} + \frac{2(\log y + \log n)^2}{y^2} \, dx \, dy + 2 \int_0^{1/n} \int_0^y 4n^2 \, dx \, dy \\ &\leq 2 \int_{1/n}^1 \int_0^y \frac{8}{y^2} + \frac{2(\log y - \log x)^2}{y^2} \, dx \, dy + 2 \int_0^{1/n} \int_0^y 4n^2 \, dx \, dy. \end{aligned}$$

Moreover, we have that

$$\int_{1/n}^1 \int_0^y \frac{8}{y^2} \, dx \, dy = \int_{1/n}^1 \frac{8}{y} \, dy = 8 \log n \quad \text{and} \quad \int_0^{1/n} \int_0^y 4n^2 \, dx \, dy = 2$$

as well as

$$\int_{1/n}^1 \int_0^y \frac{2(\log y - \log x)^2}{y^2} \, dx \, dy = \int_{1/n}^1 \frac{4}{y} \, dy = 4 \log n.$$

Combining these estimates and using that  $\frac{1}{2} \log n \leq 2 \log n - 2(1 - 1/n)$ , we obtain that, for  $n \geq 2$ ,

$$K_n \leq 2 \frac{(4 + 24 \log n)^{d/2}}{(2 \log n - 2(1 - 1/n))^d} \leq 2 \frac{(30 \log n)^{d/2}}{((\log n)/2)^d} \leq 2 \left( \frac{120}{\log n} \right)^{d/2}$$

and thus we can choose

$$\Delta_n = \frac{1}{8\sqrt{2}} \left( \frac{\log n}{120} \right)^{d/4},$$

since  $\alpha(2) = 1/2$ . The result then follows from Theorem 1 and Theorem 5.  $\square$

## 5 Application to fractional Brownian motion

We now present our second application of the results obtained in Section 3 by considering a discretized version of the quadratic variation of a fractional Brownian motion. Recall that a fractional Brownian motion  $B^H = (B_t^H : t \geq 0)$  with Hurst index  $0 < H < 1$  is a continuous-time centred Gaussian process with covariance

$$\mathbb{E}[B_t^H B_s^H] = \frac{1}{2} (t^{2H} + s^{2H} - |t - s|^{2H}), \quad s, t \geq 0.$$

If  $H = 1/2$ , then  $B^H$  is the ordinary Brownian motion, while for  $H > 1/2$  the fractional Brownian motion is a commonly used model for long-range dependencies, see [12] for details and background material. In practice, it is crucial to estimate  $H$  from given data. A well known estimator is based on the discretized quadratic variation of  $B^H$  at scale  $1/n$  on the interval  $[0, 1]$  and is defined as

$$S_n := \sum_{k=0}^{n-1} (B_{\frac{k+1}{n}}^H - B_{\frac{k}{n}}^H)^2, \quad n \geq 1.$$

From [12, Equation (2.12)] it is known that, as  $n \rightarrow \infty$ , the random variables  $n^{2H-1} S_n$ ,  $n \geq 1$ , converge in probability to 1 so that a reasonable estimator  $\hat{H}_n$  for the Hurst index  $H$  is given by

$$\hat{H}_n = \frac{1}{2} - \frac{\log S_n}{2 \log n}, \quad n \geq 2.$$

To investigate the asymptotic distributional behaviour of  $\widehat{H}_n$ , define a sequence  $(F_n : n \geq 1)$  of centred and normalized versions of the discretized quadratic variation of  $B^H$  by

$$F_n := \frac{n^{2H}}{\sigma_n} \sum_{k=0}^{n-1} [(B_{\frac{k+1}{n}}^H - B_{\frac{k}{n}}^H)^2 - n^{-2H}], \quad n \geq 1, \quad (16)$$

where we choose  $\sigma_n$  in such a way that  $\mathbb{E}[F_n^2] = 2$  (this normalization is adapted to the set-up of Section 3). Now, a short computation reveals that

$$\widehat{H}_n - H = -\frac{\log(\sigma_n F_n / n + 1)}{2 \log n},$$

which means that the behaviour of  $F_n$  controls the error of the estimator  $\widehat{H}_n$ . For this reason  $F_n$  is studied in the sequel.

We notice that  $F_n$  has the same law as

$$\frac{1}{\sigma_n} \sum_{k=0}^{n-1} [(B_{k+1}^H - B_k^H)^2 - 1] = \frac{1}{\sigma_n} \sum_{k=0}^{n-1} H_2(B_{k+1}^H - B_k^H), \quad (17)$$

where  $H_2(x) = x^2 - 1$  is the second Hermite polynomial, explaining the alternative name second Hermite power variation for  $F_n$ . Asymptotic normality for  $F_n$  together with rates of convergence for the total variation distance has been investigated in literature, see Theorem 6.3 in [12] and Section 7.4 in [15]. In fact, if  $H > 3/4$ , then the sequence  $(F_n : n \geq 1)$  does not satisfy a central limit theorem, while for  $0 < H \leq 3/4$  it holds that

$$d_{TV}(F_n, N) = \sup_{B \subset \mathbb{R} \text{ Borel set}} |\mathbb{P}(F_n \in B) - \mathbb{P}(N \in B)| \leq A_n, \quad n \geq 2,$$

where  $N$  is a centred Gaussian random variable with variance 2 and  $A_n$  is given by

$$A_n := c_H \times \begin{cases} \frac{1}{\sqrt{n}} & : 0 < H < 5/8 \\ \frac{(\log n)^{3/2}}{\sqrt{n}} & : H = 5/8 \\ \frac{1}{n^{3-4H}} & : 5/8 < H < 3/4 \\ \frac{1}{\log n} & : H = 3/4 \end{cases} \quad (18)$$

with a constant  $c_H > 0$  only depending on  $H$ . As a consequence, one can show that for  $0 < H \leq 3/4$  both of the re-scaled random variables  $\sqrt{n}(n^{2H-1}S_n - 1)$  and  $\sqrt{n} \log n (\widehat{H}_n - H)$  are, as  $n \rightarrow \infty$ , normally distributed with explicitly known limiting variances, cf. [12, Chapter 6.4].

We are now going to study the normalized versions  $F_n$  of the discretized quadratic variation functionals in more detail. For this, we will use the representation (17) for  $F_n$ . The connection between the random variables  $F_n$  and the random elements living inside a Wiener chaos of fixed order is that, for each  $n \geq 1$ ,  $F_n$  can be represented as

$$F_n = I_2^W(h_n) \quad \text{with} \quad h_n = \frac{1}{\sigma_n} \sum_{k=0}^{n-1} f_k \otimes f_k.$$

Here,  $I_2^W$  indicates the double stochastic integral with respect to a two-sided standard Brownian motion  $W$  on  $\mathbb{R}$  and  $(f_k : k \geq 1)$  is a certain sequence of square-integrable functions on  $\mathbb{R}$  such that  $I_1^W(f_k)$  has the same distribution as  $B_{k+1}^H - B_k^H$  for all  $k$  (the precise form of the  $f_k$ 's is irrelevant for our purposes). This follows from the Mandelbrot-Van Ness representation of the fractional Brownian motion as stochastic integral with respect to the ordinary two-sided Brownian motion  $W$ , see Proposition 2.3 in [12]. Our next result shows that the sequence  $(F_n : n \geq 1)$  of discretized and normalized quadratic variations of  $B^H$  satisfies a MDP as well as certain moderate and large deviation inequalities.

**Theorem 8.** Suppose that  $0 < H \leq 3/4$  and let  $F_n$  be as in (16). Then the statements in Theorem 1 and Theorem 5 hold with  $q = 2$  and  $\Delta_n = 2^{-9/4} A_n^{-1/2}$ , where  $A_n$  is given by (18).

*Proof.* Since  $q = 2$  in the language of Section 3, we have that  $K_n$  defined at (6) is just the norm of a single contraction, namely

$$K_n = \|h_n \otimes_1 h_n\|_{L^2(\mathbb{R}^2)} := \sqrt{\int_{\mathbb{R}^2} (h_n \otimes_1 h_n)(x, y)^2 \, d(x, y)},$$

where the integration is with respect to the Lebesgue measure on  $\mathbb{R}^2$ . In the proof of Theorem 6.3 in [12] it has been shown that  $\|h_n \otimes_1 h_n\|_{L^2(\mathbb{R}^2)}$  and hence  $K_n$  can be estimated from above by

$$K_n \leq \frac{1}{2\sqrt{2}} A_n$$

with  $A_n$  defined at (18). Since  $\alpha(2) = 1/2$ , the choice  $\Delta_n = 2^{-9/4} A_n^{-1/2}$  as well as Theorem 1 and Theorem 5 yield the result.  $\square$

**Remark 9.** One can more generally consider the higher-order Hermite power variations defined for  $q \geq 3$  as

$$F_n^{(q)} := \frac{1}{\sigma_n^{(q)}} \sum_{k=0}^{n-1} H_q(B_{k+1}^H - B_k^H), \quad n \geq 1,$$

where  $H_q$  is the  $q$ th Hermite polynomial and where  $\sigma_n^{(q)}$  is such that  $\mathbb{E}[F_n^{(q)}] = q!$  for all  $n \geq 1$ . These functionals can be represented as elements of the  $q$ th Wiener chaos and an estimate for  $K_n$  can in this case be deduced from Theorem 1.2 in [3] and Theorem 4.1 in [13], see also Exercise 7.5.1 in [15]. Since the results for  $q = 2$  and  $q \geq 3$  have different structures, we decided to restrict to the first case.

## 6 Application to spherical Gaussian random fields

In this section we present another application of Theorem 1 and Theorem 5 by considering the sample bispectrum of spherical random fields. These objects have recently found considerable attention especially in astrophysics, cosmology, medical imaging and geophysics, and we refer to the monograph [10] for further details on this subject. In order to simplify comparison with the existing literature, we adopt the notation from [10]. Let  $T = (T(x) : x \in \mathbb{S}^2)$  be a centred, isotropic random field on the two-dimensional unit sphere  $\mathbb{S}^2$  having finite moments up to order three. Later we additionally assume that  $T$  is Gaussian. According to [10, Theorem 5.13], for each  $x \in \mathbb{S}^2$ ,  $T(x)$  admits the harmonic representation

$$T(x) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell,m} Y_{\ell,m}(x), \quad (19)$$

where  $(Y_{\ell,m} : \ell \geq 0, -\ell \leq m \leq \ell)$  are the spherical harmonics and  $(a_{\ell,m} : \ell \geq 0, -\ell \leq m \leq \ell)$  is an array of random coefficients determined by the random field  $T$ . Abbreviating the inner sum in the above representation by  $T_\ell(x)$  it holds that  $\mathbb{E}[T_\ell(x)^2] = C_\ell \frac{2\ell+1}{4\pi}$ , independently of  $x \in \mathbb{S}^2$ , and the sequence  $(C_\ell : \ell \geq 0)$  is called the angular power spectrum of  $T$  (see Proposition 6.6 and Equation (6.21) in [10]). If  $T$  is a Gaussian random field, the angular power spectrum completely captures the dependence structure of  $T$ . In the non-Gaussian case, this structure becomes more involved and an analysis of higher-order angular power spectra is necessary. As a third-order characteristic, one can consider the integrals

$$\int_{\mathbb{S}^2} \mathbb{E}[T_{\ell_1}(x) T_{\ell_2}(x) T_{\ell_3}(x)] \, dx, \quad \ell_1, \ell_2, \ell_3 \geq 0.$$

After evaluating these expressions and re-scaling in a suitable way, one obtains the angular (average) power bispectrum  $(B_{\ell_1, \ell_2, \ell_3} : \ell_1, \ell_2, \ell_3 \geq 0)$ , which is given by

$$B_{\ell_1, \ell_2, \ell_3} := \sum_{m_1=-\ell_1}^{\ell_1} \sum_{m_2=-\ell_2}^{\ell_2} \sum_{m_3=-\ell_3}^{\ell_3} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \mathbb{E}[a_{\ell_1, m_1} a_{\ell_2, m_2} a_{\ell_3, m_3}], \quad \ell_1, \ell_2, \ell_3 \geq 0.$$

Here  $\begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix}$  is a combinatorial coefficient only depending on  $\ell_1, \ell_2, \ell_3$  and  $m_1, m_2, m_3$ , the so-called Wigner  $3j$ -coefficient for which we refer to [10, Chapter 3.5.3]. We remark that these coefficients are closely related to the Clebsch-Gordan coefficients, a commonly used tool in the representation theory of compact Lie groups. The Wigner  $3j$ -coefficients vanish unless  $m_1 + m_2 + m_3 = 0$  and  $|\ell_i - \ell_j| \leq \ell_k \leq \ell_i + \ell_j$  for all  $i, j, k \in \{1, 2, 3\}$  (see [10, Proposition 3.44]).

In the following let  $\ell_1, \ell_2, \ell_3 \geq 0$  be such that  $\ell_1 + \ell_2 + \ell_3$  is even and  $|\ell_i - \ell_j| \leq \ell_k \leq \ell_i + \ell_j$  for all  $i, j, k \in \{1, 2, 3\}$ . A high-frequency, unbiased, minimum mean square error estimator for the angular power bispectrum is given by

$$\hat{B}_{\ell_1, \ell_2, \ell_3} := \sum_{m_1=-\ell_1}^{\ell_1} \sum_{m_2=-\ell_2}^{\ell_2} \sum_{m_3=-\ell_3}^{\ell_3} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix} a_{\ell_1, m_1} a_{\ell_2, m_2} a_{\ell_3, m_3}, \quad (20)$$

where  $a_{\ell_1, m_1}, a_{\ell_2, m_2}, a_{\ell_3, m_3}$  are the observed coefficients in the harmonic representation (19) of the given realization of  $T$ . We also define its normalized version  $S_{\ell_1, \ell_2, \ell_3} := \hat{B}_{\ell_1, \ell_2, \ell_3} / \sqrt{C_{\ell_1} C_{\ell_2} C_{\ell_3}}$ , cf. [10, Chapter 9.2.2]. The estimator  $\hat{B}_{\ell_1, \ell_2, \ell_3}$  is known as the sample bispectrum of  $T$ , and we refer to  $S_{\ell_1, \ell_2, \ell_3}$  as the re-scaled sample bispectrum. Because of the symmetry we assume without loss of generality that  $\ell_1 \leq \ell_2 \leq \ell_3$ .

An important problem in the statistical investigation of spherical random fields is to test for (non-) Gaussianity. Since the sample bispectrum is a prominent test statistic, we assume from now on that  $T$  is Gaussian and study the behaviour of  $S_{\ell_1, \ell_2, \ell_3}$  under this assumption.

By Lemma 9.6 in [10], the re-scaled sample bispectrum is an element of the third Wiener chaos associated with the underlying Gaussian random field  $T$ , which we can assume to be generated by a standard Brownian motion on the unit interval. In particular,  $\mathbb{E}[S_{\ell_1, \ell_2, \ell_3}] = 0$  and  $\mathbb{E}[S_{\ell_1, \ell_2, \ell_3}^2] = D_{\ell_1, \ell_2, \ell_3}$  with  $D_{\ell_1, \ell_2, \ell_3} := 1 + \mathbf{1}\{\ell_1 = \ell_2\} + \mathbf{1}\{\ell_2 = \ell_3\} + 3\mathbf{1}\{\ell_1 = \ell_3\}$  (see [10, Theorem 9.7]). For our asymptotic investigations let  $(u_n : n \geq 1)$  and  $(v_n : n \geq 1)$  be non-negative integer-valued sequences such that  $n \leq u_n \leq v_n \leq 2n$  and suppose that  $n + u_n + v_n$  is even for all  $n \geq 1$ . We recall from [10, Theorem 9.9] that, as  $n \rightarrow \infty$ , the normalized sample bispectrum of  $T$  is asymptotically normal and that

$$d_{TV} \left( \sqrt{\frac{6}{D_{n, u_n, v_n}}} S_{n, u_n, v_n}, N \right) \leq \sqrt{\frac{32}{3n}}, \quad n \geq 1,$$

where  $N$  is a Gaussian random variable with variance  $3! = 6$ . Using the theory developed in Section 3, we can add a moderate deviation principle as well as certain moderate and large deviation estimates for the normalized sample bispectrum.

**Theorem 10.** *Let  $(u_n : n \geq 1)$  and  $(v_n : n \geq 1)$  be non-negative integer-valued sequences such that  $n \leq u_n \leq v_n \leq 2n$  and  $n + u_n + v_n$  is even for all  $n \geq 1$  and let  $F_n := \sqrt{\frac{6}{D_{n, u_n, v_n}}} S_{n, u_n, v_n}$ . Then the statements of Theorem 1 and Theorem 5 hold with  $q = 3$  and  $\Delta_n = 3^{-9/2} (\sqrt{3n}/2)^{5/12}$ ,  $n \geq 1$ .*

*Proof.* Since the normalized sample bispectrum is an element of the third Wiener chaos, there exists a sequence  $(h_{n, u_n, v_n} : n \geq 1)$  of square-integrable and symmetric functions on  $[0, 1]^3$  (supplied with the Lebesgue measure) such that  $\|h_{n, u_n, v_n}\|_{L^2([0, 1]^3)} = 1$  and

$$\sqrt{\frac{6}{D_{n, u_n, v_n}}} S_{n, u_n, v_n} = I_3^W(h_{n, u_n, v_n})$$

for all  $n \geq 1$ , where  $I_3^W$  stands for a multiple stochastic integral of order three with respect to a standard Brownian motion  $W$  on  $[0, 1]$ . It has been shown in the proof of Theorem 9.7 in [10] that the fourth cumulant  $\text{cum}_4(I_3^W(h_{n,u_n,v_n}))$  of  $I_3^W(h_{n,u_n,v_n})$  is bounded from above by  $432/n$ . Consequently, recalling the definition (8), we find that

$$L_n = \frac{\sqrt{\text{cum}_4(I_3^W(h_{n,u_n,v_n}))}}{3 \cdot 3!} \leq \frac{1}{3 \cdot 3!} \sqrt{\frac{432}{n}} = \frac{2}{\sqrt{3n}},$$

and hence the conditions of Corollary 2 and Theorem 5 (iii) are satisfied with  $q = 3$  and

$$\Delta_n = 3^{-9/2} \left( \frac{\sqrt{3n}}{2} \right)^{5/12},$$

since  $\alpha(3) = 5/12$ . This proves the claim.  $\square$

## 7 Proofs of the main results

The next lemma is our main device to prove Theorem 1 and Theorem 5. It summarizes a moderate deviation principle and fine probability estimates, which are available under certain bounds on cumulants. This approach goes back to the ‘Lithuanian school of probability’, and we refer especially to the monograph [21].

**Lemma 11.** *Let  $(X_n : n \geq 1)$  be a sequence of real-valued random variables such that  $\mathbb{E}[X_n] = 0$ ,  $\mathbb{E}[X_n^2] = \sigma^2 \geq 1$  and  $\mathbb{E}[|X_n|^m] < \infty$  for all  $m \geq 1$ . Suppose that there is a constant  $\gamma \geq 0$  such that the cumulants of  $X_n$  satisfy*

$$|\text{cum}_m(X_n)| \leq \frac{(m!)^{1+\gamma}}{\Delta_n^{m-2}} \quad \text{for all } m \geq 3 \quad (21)$$

with  $\Delta_n > 0$  for  $n \geq 1$ .

(a) *Let  $(a_n : n \geq 1)$  be a real sequence such that*

$$\lim_{n \rightarrow \infty} a_n = \infty \quad \text{and} \quad \lim_{n \rightarrow \infty} a_n / \Delta_n^{1/(1+2\gamma)} = 0.$$

*Then the sequence of re-scaled random variables  $(a_n^{-1} X_n : n \geq 1)$  satisfies a MDP with speed  $a_n^2$  and Gaussian rate function  $\mathcal{I}(z) = z^2/(2\sigma^2)$ .*

(b) *There exist constants  $c_0, c_1, c_2 > 0$  only depending on  $\gamma$  such that for  $\Delta_n \geq c_0$  and  $0 \leq z \leq c_1 \sigma \Delta_n^{1/(1+2\gamma)}$ ,*

$$\left| \log \frac{\mathbb{P}(X_n \geq z)}{1 - \Phi_{\sigma^2}(z)} \right| \leq c_2 \frac{1 + (z/\sigma)^3}{\Delta_n^{1/(1+2\gamma)}} \quad \text{and} \quad \left| \log \frac{\mathbb{P}(X_n \leq -z)}{\Phi_{\sigma^2}(-z)} \right| \leq c_2 \frac{1 + (z/\sigma)^3}{\Delta_n^{1/(1+2\gamma)}},$$

*where  $\Phi_{\sigma^2}$  is the distribution function of a centred Gaussian random variable with variance  $\sigma^2$ .*

(c) *One has that*

$$\mathbb{P}(|X_n| \geq z) \leq 2 \exp \left( -\frac{1}{4} \min \left\{ \frac{z^2}{2^{1+\gamma}}, (z \Delta_n)^{1/(1+\gamma)} \right\} \right)$$

*for all  $z \geq 0$ .*

*Proof.* Define  $\tilde{X}_n := X_n/\sigma$  and observe that by  $\sigma^2 \geq 1$  and (21),

$$|\text{cum}_m(\tilde{X}_n)| = \sigma^{-m} |\text{cum}_m(X_n)| \leq |\text{cum}_m(X_n)| \leq \frac{(m!)^{1+\gamma}}{\Delta_n^{m-2}}.$$

Applying [5, Theorem 1.1] and [6, Corollary 3.2] to  $(\tilde{X}_n : n \geq 1)$  yields part (a) and part (b). The assertion in (c) follows from the Corollary after Lemma 2.4 in [21] with  $H = 2^{1+\gamma}$  there. As discussed in Remark 6 above, the results in (b) and (c) are simplified versions of the ‘main lemmas’ from [20] and [1], respectively, which are summarized in Chapter 2 of [21].  $\square$

**Remark 12.** (i) To require the estimate (21) is a rather natural condition from the viewpoint of complex analysis. As discussed at the beginning of Chapter 2 in [21], the cumulant bound (21) with  $\gamma = 0$  implies analyticity of the cumulant generating functions  $\log \mathbb{E}[\exp(wX_n)]$ ,  $w \in \mathbb{C}$ , of the random variables  $(X_n : n \geq 1)$  in the discs  $\{|w| \leq \Delta_n\} \subset \mathbb{C}$ . On the other hand, if the cumulant generating function is analytic in a disc around the origin with radius  $\Delta_n$  and such that

$$\sup_{|w|=\Delta_n} |\log \mathbb{E}[\exp(wX_n)]| \leq \Delta_n^2,$$

then the random variable  $X_n$  satisfies the cumulant bound (21) by Cauchy’s integral formula for derivatives (recall (5)).

To allow for  $\gamma > 0$  takes into account a heavy tail behaviour and a super-exponential growth of cumulants (or moments) of the involved random variables. For example, if a random variable has density  $x \mapsto \frac{\alpha}{2}|x|^{\alpha-1}e^{-|x|^\alpha}$  for some  $\alpha > 0$ , then the  $m$ th moment is  $\Gamma(1 + m/\alpha)$  if  $m$  is even and zero otherwise. Thus, Stirling’s formula together with [21, Lemma 3.1] implies that the cumulant bound (21) is satisfied with  $\gamma = \frac{1}{\alpha} - 1$  if  $0 < \alpha \leq 1$  and  $\gamma = 0$  if  $\alpha > 1$ . In view of the tail estimate (13) it is therefore not unexpected that we have  $\gamma = \frac{q}{2} - 1$  for random variables belonging to the  $q$ th Wiener chaos  $\mathcal{C}_q$ ,  $q \geq 2$ . We also refer to Remark 19 below for further discussion of this point.

(ii) Under condition (21) one also has the Berry-Esseen estimate

$$\sup_{x \in \mathbb{R}} |\mathbb{P}(X_n \leq x) - \Phi_{\sigma^2}(x)| \leq c \Delta_n^{-1/(1+2\gamma)} \quad (22)$$

with a constant  $c > 0$  only depending on  $\gamma$ , see [21, Corollary 2.1]. In the context of Theorem 1 this leads to a rate of convergence of order  $K_n^{\alpha(q)/(q-1)}$ . To the best of our knowledge, this provides a first proof of the fourth moment theorem including rates of convergence without resorting to Stein’s method. However, comparing this with the bound (2) derived via the Malliavin-Stein method in [13] (see also [15, Chapter 5.2]), we see that the rate of convergence via the method of cumulants is weaker for all  $q \geq 2$ . Moreover, the bound in (2) is for the total variation distance, which is larger than the left-hand side of (22). For this reason, we do not pursue rates for the normal approximation further in this paper.

Our strategy for the proof of Theorem 1 and Theorem 5 is to establish for the random variable  $F_n = I_q(h_n)$  the cumulant bound (21). In what follows we assume without loss of generality that  $\mathfrak{H} = L^2(A, \mathcal{A}, \mu) =: L^2(A)$  with a Polish space  $(A, \mathcal{A})$  and a non-atomic  $\sigma$ -finite measure  $\mu$ . This is possible because of isomorphy of Hilbert spaces. Recall that we denote by  $L_s^2(A^n)$ ,  $n \geq 1$ , the subspace of  $L^2(A^n)$  consisting of symmetric functions, i.e., functions which are invariant under permutation of their arguments. Moreover, the tensor product of two functions  $f_1 : A^{n_1} \rightarrow \mathbb{R}$  and  $f_2 : A^{n_2} \rightarrow \mathbb{R}$ ,  $n_1, n_2 \geq 1$ , is a function  $f_1 \otimes f_2 : A^{n_1+n_2} \rightarrow \mathbb{R}$  given by

$$f_1 \otimes f_2(x_1, \dots, x_{n_1+n_2}) = f_1(x_1, \dots, x_{n_1})f_2(x_{n_1+1}, \dots, x_{n_1+n_2}).$$

For integers  $\ell \geq 1$  and  $n_1, \dots, n_\ell \geq 1$  define  $N_0 := 0$  and, for  $i \in \{1, \dots, \ell\}$ , put  $N_i := n_1 + \dots + n_i$  and  $J_i := \{N_{i-1} + 1, \dots, N_i\}$ . By a partition  $\sigma$  of  $\{1, \dots, N_\ell\}$  we understand a collection  $\{B_1, \dots, B_{|\sigma|}\}$  of non-empty and pairwise disjoint subsets  $B_1, \dots, B_{|\sigma|} \subset \{1, \dots, N_\ell\}$  with  $B_1 \cup \dots \cup B_{|\sigma|} = \{1, \dots, N_\ell\}$ . The sets  $B_1, \dots, B_{|\sigma|}$  are called blocks, and by  $|\sigma|$  we denote the number of blocks of  $\sigma$ . Let  $\Pi(n_1, \dots, n_\ell)$  be the set of partitions  $\sigma$  of  $\{1, \dots, N_\ell\}$  satisfying

- $|B \cap J_i| \leq 1$  for all  $i \in \{1, \dots, \ell\}$  and blocks  $B \in \sigma$ ,
- $|B| = 2$  for all blocks  $B$  of  $\sigma$ ,
- for all non-empty sets  $M_1, M_2 \subset \{1, \dots, \ell\}$  with  $M_1 \cup M_2 = \{1, \dots, \ell\}$  there are a block  $B \in \sigma$  and elements  $i_1 \in M_1$  and  $i_2 \in M_2$  such that  $B \cap J_{i_1} \neq \emptyset$  and  $B \cap J_{i_2} \neq \emptyset$ .

For brevity, we also write  $\Pi(q[m], q_1, \dots, q_k)$  instead of  $\Pi(q, \dots, q, q_1, \dots, q_k)$ , where  $q$  appears  $m$  times. Note that  $\Pi(n_1, \dots, n_\ell)$  can be empty, in particular if  $N_\ell$  is odd.

For a function  $f : A^{N_\ell} \rightarrow \mathbb{R}$  and a partition  $\sigma \in \Pi(n_1, \dots, n_\ell)$  we define  $f_\sigma : A^{|\sigma|} \rightarrow \mathbb{R}$  by replacing all arguments of  $f$  with indices belonging to the same block of  $\sigma$  by a new common variable. For example, if  $f : A^4 \rightarrow \mathbb{R}$  and  $\sigma = \{\{1, 4\}, \{2, 3\}\}$ , we have that  $f_\sigma(y, z) = f(y, z, z, y)$ . This notation allows us to recall from [18, Corollary 7.3.1] (see also Proposition 5.6 in [14]) the following classical expression for the cumulants of a random variable of the type  $I_q(h)$  with  $h \in L_s^2(A^q)$ .

**Lemma 13.** *For  $q \geq 1$  and  $h \in L_s^2(A^q)$ ,*

$$\text{cum}_m(I_q(h)) = \sum_{\sigma \in \Pi(q[m])} \int_{A^{|\sigma|}} (h^{\otimes m})_\sigma d\mu^{|\sigma|}, \quad m \geq 1,$$

where the sum on the right-hand side has to be interpreted as 0 if  $\Pi(q[m]) = \emptyset$ .

A crucial step in our proof of Theorem 1 and Theorem 5 is to re-write the right-hand side of the cumulant expression provided in Lemma 13 in terms of contractions. The underlying idea is to consider for a fixed partition  $\sigma \in \Pi(q[m])$  the functions in the tensor product  $h^{\otimes m}$  appearing on the right-hand side in Lemma 13 as vertices of a suitable multigraph induced by  $\sigma$  and to group them in an appropriate way according to a maximal matching.

More precisely, with  $\sigma \in \Pi(q[m])$  we associate a multigraph  $G_\sigma$  as follows. The set of vertices is  $\{1, \dots, m\}$ , and for each block  $B$  of  $\sigma$  with  $B = (B \cap J_i) \cup (B \cap J_j)$  we connect the vertices  $i$  and  $j$  by an edge. In particular, each vertex of  $G_\sigma$  has degree  $q$  (in other words this means that  $G_\sigma$  is  $q$ -regular), and  $G_\sigma$  is a connected multigraph without loops.

A matching of  $G_\sigma$  is a set of non-adjacent edges of  $G_\sigma$ , and we denote by  $M(G_\sigma)$  the maximal size of such a set. This so-called matching number of  $G_\sigma$  is an important quantity considered in combinatorics. For the matching number of a multigraph  $G_\sigma$  associated with a partition  $\sigma \in \Pi(q[m])$  we have the following lower bound.

**Lemma 14.** *For  $q \geq 2$  and  $m \geq 3$  suppose that  $\Pi(q[m]) \neq \emptyset$  and fix  $\sigma \in \Pi(q[m])$ . Then*

$$M(G_\sigma) \geq L(q, m) := \begin{cases} \left\lceil \frac{(q^2 - q - 1)m - (q - 1)}{q(3q - 5)} \right\rceil & : q \text{ odd} \\ \min \left( \left\lfloor \frac{m}{2} \right\rfloor, \left\lceil \frac{(q + 2)m}{3q + 2} \right\rceil \right) & : q \text{ even}. \end{cases} \quad (23)$$

*Proof.* From Theorem 1 (with  $\lambda = 2$  there) and Theorem 2 in [11] it follows that  $M(G_\sigma)$  satisfies the inequality  $M(G_\sigma) \geq L(q, m)$  for  $q \geq 3$ . If  $q = 2$ ,  $G_\sigma$  is a cycle and hence also satisfies the estimate (23).  $\square$

The construction of the multigraph  $G_\sigma$  described above and the lower bound on the matching number in Lemma 14 allow us to re-write the summands on the right-hand side of the cumulant expression in Lemma 13 in the following way:

**Lemma 15.** For  $q \geq 2$  and  $m \geq 3$  suppose that  $\Pi(q[m]) \neq \emptyset$  and fix  $\sigma \in \Pi(q[m])$  and  $h \in L_s^2(A^q)$ . Then there are non-negative integers  $m_1, m_2$  satisfying  $m_1 + 2m_2 = m$  and  $L(q, m) \leq m_2 \leq m/2$  and  $r_i \in \{1, \dots, q-1\}$  for  $i \in \{1, \dots, m_2\}$  as well as a partition  $\tilde{\sigma} \in \Pi(q[m_1], q - r_1, \dots, q - r_{m_2})$  such that

$$\int_{A^{|\sigma|}} (h^{\otimes m})_{\sigma} d\mu^{|\sigma|} = \int_{A^{|\tilde{\sigma}|}} \left( h^{\otimes m_1} \otimes \bigotimes_{i=1}^{m_2} (h \otimes_{r_i} h) \right)_{\tilde{\sigma}} d\mu^{|\tilde{\sigma}|}.$$

*Proof.* We construct the multigraph  $G_{\sigma}$  as described above and choose a matching of maximal cardinality. Then we split the set of variables of  $(h^{\otimes m})_{\sigma}$  into two groups, those variables belonging to vertices (factors in the tensor product  $h^{\otimes m}$ ) that are matched and the remaining variables. Now, we apply Fubini's theorem and integrate over the first group of variables. By this construction,  $m_2 = M(G_{\sigma})$  pairs of functions  $h$  are transformed into terms of the type  $h \otimes_{r_i} h$  with  $r_i \in \{1, \dots, q-1\}$ ,  $i = 1, \dots, m_2$ . In fact, that  $r_i < q$  follows since  $G_{\sigma}$  is connected by construction. We write the integration with respect to the variables belonging to the second group in terms of a partition  $\tilde{\sigma} \in \Pi(q[m - 2M(G_{\sigma})], q - r_1, \dots, q - r_{m_2})$  and set  $m_1 := m - 2M(G_{\sigma})$ . Now the observation from Lemma 14 that  $m_2 = M(G_{\sigma}) \geq L(q, m)$  concludes the proof.  $\square$

**Remark 16.** It follows from the examples in Section 3 in [11] that for sufficiently large  $m$  there are partitions  $\sigma$  such that inequality (23) is sharp in that  $M(G_{\sigma}) = L(q, m)$ , and hence  $m_2 = L(q, m)$  and  $m_1 = m - 2L(q, m)$  in Lemma 15.

Finally, let us recall the following generalized Cauchy-Schwarz inequality from Lemma 4.1 in [2].

**Lemma 17.** Fix  $\ell \geq 2$ ,  $n_1, \dots, n_{\ell} \geq 1$  such that  $\Pi(n_1, \dots, n_{\ell}) \neq \emptyset$ ,  $f_i \in L_s^2(A^{n_i})$  for  $i \in \{1, \dots, \ell\}$  and  $\sigma \in \Pi(n_1, \dots, n_{\ell})$ . Then

$$\int_{A^{|\sigma|}} \left( \bigotimes_{i=1}^{\ell} |f_i| \right)_{\sigma} d\mu^{|\sigma|} \leq \prod_{i=1}^{\ell} \|f_i\|_{L^2(A^{n_i})}.$$

**Remark 18.** If  $B \in \mathcal{A}$  with  $\mu(B) < \infty$  and each of the functions  $f_i$  is of the form

$$f_i(x_1, \dots, x_{n_i}) = \prod_{j=1}^{n_i} \mathbf{1}(x_j \in B),$$

we have equality in Lemma 17.

After these preparations, we can now establish the cumulant bound (21) for the random variable  $F_n = I_q(h_n)$ .

*Proof of Theorem 1 and Theorem 5.* Using the cumulant formula provided in Lemma 13 we have that

$$\text{cum}_m(F_n) = \sum_{\sigma \in \Pi(q[m])} \int_{A^{|\sigma|}} (h_n^{\otimes m})_{\sigma} d\mu^{|\sigma|}.$$

To each summand we apply Lemma 15 with the notation introduced there and Lemma 17 to see that

$$\begin{aligned} \left| \int_{A^{|\sigma|}} (h_n^{\otimes m})_{\sigma} d\mu^{|\sigma|} \right| &\leq \int_{A^{|\tilde{\sigma}|}} \left( |h_n|^{\otimes m_1} \otimes \bigotimes_{i=1}^{m_2} |h_n \otimes_{r_i} h_n| \right)_{\tilde{\sigma}} d\mu^{|\tilde{\sigma}|} \\ &\leq \|h_n\|_{L^2(A^q)}^{m_1} \prod_{i=1}^{m_2} \|h_n \otimes_{r_i} h_n\|_{L^2(A^{2(q-r_i)})} \\ &\leq K_n^{L(q, m)} \end{aligned} \tag{24}$$



with  $K_n$  defined at (6). Here, we have used that  $L(q, m) \leq m_2$ , the assumption that  $\|h_n\|_{L^2(A^q)} = 1$  and that  $K_n \leq 1$ . The latter property is a consequence of the Cauchy-Schwarz inequality, implying that  $\|h_n \otimes_r h_n\|_{L^2(A^{2(q-r)})}^2 \leq \|h_n\|_{L^2(A^q)}^4 = 1$  for all  $r \in \{1, \dots, q-1\}$ . Thus, (24) yields that

$$|\text{cum}_m(F_n)| \leq |\Pi(q[m])| K_n^{L(q,m)}.$$

It has been shown in Proposition 5.3 of [21] that  $|\Pi(q[m])|$  is bounded from above by

$$|\Pi(q[m])| \leq (m!)^{q/2} (q^{q/2})^m. \quad (25)$$

Moreover, it follows from the definition of  $L(q, m)$  that  $L(q, m) \geq \alpha(q)(m-2)$  with  $\alpha(q)$  defined at (7). Consequently,

$$|\text{cum}_m(F_n)| \leq (m!)^{q/2} (q^{q/2})^m K_n^{\alpha(q)(m-2)} \leq (m!)^{q/2} (q^{3q/2})^{m-2} K_n^{\alpha(q)(m-2)},$$

where we have used that  $3(m-2) \geq m$  for  $m \geq 3$ . Choosing

$$\gamma = \frac{q}{2} - 1 \quad \text{and} \quad \Delta_n^{-1} = q^{3q/2} K_n^{\alpha(q)}$$

establishes the cumulant bound (21). In view of Lemma 11 this concludes the proof of Theorem 1 and Theorem 5.  $\square$

**Remark 19.** (i) Proposition 5.3 in [21] also gives the lower bound

$$|\Pi(q[m])| \geq \frac{1}{8} (m!)^{q/2} (\sqrt{2})^m. \quad (26)$$

Comparison of (25) and (26) shows that we can in general choose  $\gamma$  not smaller than  $\frac{q}{2} - 1$ . This goes hand in hand with the observations made in Remark 12 above.

In contrast to our situation, it is typically a difficult task to decide whether for given random variables the parameter  $\gamma$  in the cumulant estimate (21) is optimal or not. Such a situation arises, for example, in [6], where problems from geometric probability have been considered. There,  $\gamma$  is different from zero and depends on the particular model and even on the space dimension.

- (ii) Because of the sharpness of the lower bound on the matching number and the sharpness of the generalized Cauchy-Schwarz inequality discussed in Remarks 16 and 18, there are situations for which one has equality in all estimates of (24). This shows that the exponent of  $K_n$  in  $\Delta_n$  is optimal in general.

We finally establish Proposition 3 based on arguments from the proof of Theorem 6.12 in [7].

*Proof of Proposition 3.* Since  $F_n$  is a linear combination of  $H_q(X(e_k))$ ,  $k \in \{1, \dots, n\}$ , it is by definition an element of the  $q$ th Wiener chaos  $\mathcal{C}_q$ , and we have that  $F_n = I_q(h_n)$  with  $h_n = \frac{1}{\sqrt{n}} \sum_{k=1}^n e_k^{\otimes q}$  for  $n \geq 1$ . Hence,

$$\text{Var } F_n = \mathbb{E}[I_q(h_n)^2] = q! \|h_n\|_{\mathfrak{H}^{\otimes q}}^2 = q!$$

and

$$K_n = \max_{r=1, \dots, q-1} \|h_n \otimes_r h_n\|_{\mathfrak{H}^{\otimes 2(q-r)}} = \max_{r=1, \dots, q-1} \frac{1}{n} \left\| \sum_{k=1}^n e_k^{\otimes 2(q-r)} \right\|_{\mathfrak{H}^{\otimes 2(q-r)}} = \frac{1}{\sqrt{n}}.$$

For  $n \geq 1$  define  $S_n = \sum_{k=1}^n H_q(X(e_k))$  and note that  $S_2$  can be regarded as a polynomial of degree  $q$  depending on the random variables  $X(e_1)$  and  $X(e_2)$ . Thus, it follows from Equation (6.10) in [7] that there are constants  $\tilde{C}, \tilde{c}, \tilde{t}_0 > 0$  such that

$$\mathbb{P}(|S_2| \geq t) \geq \tilde{C} \exp(-\tilde{c}t^{2/q}) \quad \text{for all } t \geq \tilde{t}_0. \quad (27)$$

Since  $(S_n : n \geq 1)$  is a martingale with respect to the natural filtration induced by the random variables  $(X(e_k) : k \geq 1)$  we have that

$$\mathbb{E}[|S_n| \mid X(e_1), X(e_2)] \geq |S_2| \quad \mathbb{P}\text{-a.s.} \quad (28)$$

for  $n \geq 2$ . Hence,

$$\mathbb{P}(|S_n| \geq u) \geq \mathbb{P}(|S_n| \geq |S_2|/2, |S_2| \geq 2u) \geq \mathbb{P}(|S_n| \geq \mathbb{E}[|S_n| \mid X(e_1), X(e_2)]/2, |S_2| \geq 2u),$$

where we used (28) for the second inequality. After re-writing, we get

$$\mathbb{P}(|S_n| \geq u) \geq \mathbb{E}[\mathbb{E}[\mathbf{1}\{|S_n| \geq \mathbb{E}[|S_n| \mid X(e_1), X(e_2)]/2\} \mid X(e_1), X(e_2)] \mathbf{1}\{|S_2| \geq 2u\}].$$

If  $X(e_1)$  and  $X(e_2)$  are given,  $S_n$  is the sum of an element in the  $q$ th Wiener chaos and a constant so that Theorem 6.9 in [7] yields that

$$\mathbb{E}[\mathbf{1}\{|S_n| \geq \mathbb{E}[|S_n| \mid X(e_1), X(e_2)]/2\} \mid X(e_1), X(e_2)] \geq c_q \quad \mathbb{P}\text{-a.s.}$$

with a constant  $c_q > 0$  only depending on  $q$ . We thus obtain that

$$\mathbb{P}(|S_n| \geq u) \geq c_q \mathbb{P}(|S_2| \geq 2u).$$

Now, the estimate (27) yields, for  $z \geq z_0 := \tilde{t}_0$ , that

$$\mathbb{P}(|F_n| \geq z) = \mathbb{P}(|S_n| \geq \sqrt{n}z) \geq c_q \tilde{C} \exp(-2^{2/q} \tilde{c} n^{1/q} z^{2/q}),$$

which proves (9).

Since  $a_n \rightarrow \infty$ , as  $n \rightarrow \infty$ , it follows from (9) that, for any  $t > 0$ ,

$$\begin{aligned} \limsup_{n \rightarrow \infty} a_n^{-2} \log \mathbb{P}(a_n^{-1} F_n \geq t) &\geq \limsup_{n \rightarrow \infty} a_n^{-2} \log(C \exp(-c n^{1/q} (a_n t)^{2/q})) \\ &= \limsup_{n \rightarrow \infty} \frac{-c n^{1/q} (a_n t)^{2/q}}{a_n^2}. \end{aligned}$$

If  $n^{1/(2q-2)}/a_n \rightarrow 0$  as  $n \rightarrow \infty$ , the right-hand side converges to zero, implying that  $(a_n^{-1} F_n : n \geq 1)$  cannot satisfy a MDP with speed  $a_n^2$  and Gaussian rate function.  $\square$

## Acknowledgements

Parts of this paper were written during a Research-in-Pairs stay of the authors at Mathematisches Forschungsinstitut Oberwolfach. All support is gratefully acknowledged. We also thank Sabine Jansen for a stimulating discussion.

MS has been funded by the German Research Foundation (DFG) through the research unit ‘‘Geometry and Physics of Spatial Random Systems’’ under the grant HU 1874/3-1. CT has been supported by the German research foundation (DFG) via SFB-TR 12.

## References

- [1] V. BENTKUS AND R. RUDZKIS: *Exponential estimates of the distribution of random variables*, Lithuanian Math. J. **20**, 15–30 (1980).
- [2] H. BIERMÉ, A. BONAMI, I. NOURDIN AND G. PECCATI: *Optimal Berry-Esseen rates on the Wiener space: the barrier of third and fourth cumulants*, ALEA Lat. Am. J. Probab. Math. Stat. **9**, 473–500 (2012).
- [3] J.-C. BRETON AND I. NOURDIN: *Error bounds on the non-normal approximation of Hermite power variations of fractional Brownian motion*, Electron. Comm. Probab. **13**, 482–493 (2008).
- [4] A. DEMBO AND O. ZEITOUNI: *Large Deviations. Techniques and Applications*, 2nd Edition, Springer, New York (1998).
- [5] H. DÖRING AND P. EICHELSBACHER: *Moderate deviations via cumulants*, J. Theor. Probab. **26**, 360–385 (2013).
- [6] P. EICHELSBACHER, T. SCHREIBER AND M. RAIČ: *Moderate deviations for stabilizing functionals in geometric probability*, to appear in Ann. Inst. H. Poincaré Probab. Statist. (2014).
- [7] S. JANSON: *Gaussian Hilbert Spaces*, Cambridge University Press, Cambridge (1997).
- [8] T. JEULIN: *Semimartingales et Groissement d’une Filtration*, Lecture Notes in Mathematics **833**, Springer, Berlin (1980).
- [9] P. MAJOR: *On the estimation of multiple random integrals and degenerate U-statistics*, Lecture Notes in Mathematics **2079**, Springer, Berlin (2013).
- [10] D. MARINUCCI AND G. PECCATI: *Random Fields on the Sphere*, Cambridge University Press, Cambridge (2011).
- [11] T. NISHIZEKI: *On the maximum matchings of regular multigraphs*, Discrete Math. **37**, 105–114 (1981).
- [12] I. NOURDIN: *Selected Aspects of Fractional Brownian Motion*, Bocconi University Press and Springer, Milan (2012).
- [13] I. NOURDIN AND G. PECCATI: *Stein’s method on Wiener chaos*, Probab. Theory Related Fields **145**, 75–118 (2009).
- [14] I. NOURDIN AND G. PECCATI: *Cumulants on the Wiener space*, J. Funct. Anal. **258**, 3775–3791 (2010).
- [15] I. NOURDIN AND G. PECCATI: *Normal Approximations with Malliavin Calculus: From Stein’s Method to Universality*, Cambridge University Press, Cambridge (2012).
- [16] I. NOURDIN AND G. PECCATI: *The optimal fourth moment theorem*, to appear in Proc. Am. Math. Soc. (2014).
- [17] D. NUALART AND G. PECCATI: *Central limit theorems for sequences of multiple stochastic integrals*, Ann. Probab. **33**, 177–193 (2005).
- [18] G. PECCATI AND M.S. TAQQU: *Wiener Chaos: Moments, Cumulants and Diagrams*, Bocconi University Press and Springer, Milan (2011).
- [19] G. PECCATI AND M. YOR: *Hardy’s inequality in  $L^2([0,1])$  and principal values of Brownian local times*, In: Asymptotic Methods in Stochastics (Eds. L. Horváth and B. Szyszkowicz), Fields Institute Communications Series, American Mathematical Society, Providence (2004).
- [20] R. RUDZKIS, L. SAULIS AND V.A. STATULEVIČIUS: *A general lemma on probabilities of large deviations*, Lithuanian Math. J. **18**, 99–116 (1978).
- [21] L. SAULIS AND V.A. STATULEVIČIUS: *Limit Theorems for Large Deviations*, Kluwer Academic Publishers, Dordrecht (1991).